This Raw Listing contains the General Information Section and up to the first 5 pages.

ENTERED

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: RADIN, DAVID N.
CRAMER, CAROLE L.
OISHI, KAREN K.
WEISSENBORN, DEBORAH L.

(ii) TITLE OF INVENTION: PRODUCTION OF LYSOSOMAL ENZYMES IN PLANT-BASED EXPRESSION SYSTEMS

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Pennie & Edmonds
(B) STREET: 1155 Avenue of the Americas
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: USA
(F) ZIP: 10036-2711

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08/713,928
(B) FILING DATE: 13-SEP-1996
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/003,737
(B) FILING DATE: 14-SEP-1995

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Coruzzi, Laura A.
(B) REGISTRATION NUMBER: 30,742
(C) REFERENCE/DOCKET NUMBER: 7956-0011-999

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 790-9090
(B) TELEFAX: (212) 869-9741
(C) TELEX: 66141 PENNIE

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTGTCTAGAG TAAGCATCAT GGCTGGC 27

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CACGAATTCT GGCGACGCCA CAGGTAGGTG TGA 33

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1642 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
100  ATGGAGTTTT CAAGTCCTTC CAGAGAGGAA TGTCCCAAGC CTTTGAGTAG GGTAAGCATC      60
101
102  ATGGCTGGCA GCCTCACAGG TTTGCTTCTA CTTCAGGCAG TGTCGTGGGC ATCAGGTGCC     120
103
104  CGCCCCTGCA TCCCTAAAAG CTTCGGCTAC AGCTCGGTGG TGTGTGTCTG CAATGCCACA     180
105
106  TACTGTGACT CCTTTGACCC CCCGACCTTT CCTGCCCTTG GTACCTTCAG CCGCTATGAG     240
107
108  AGTACACGCA GTGGGCGACG GATGGGGCTG AGTATGGGGC CCATCCAGGC TAATCACACG     300
109
110  GGCACAGGCC TGCTACTGAC CCTGCAGCCA GAACAGAAGT TCCAGAAAGT GAAGGGATTT     360
111
112  GGAGGGGCCA TGACAGATGC TGCTGCTCTC AACATCCTTG CCCTGTCACC CCCTGCCCAA     420
113
114  AATTTGCTAC TTAAATCGTA CTTCTCTGAA GAAGGAATCG GATATAACAT CATCCGGGTA     480
115
116  CCCATGGCCA GCTGTGACTT CTCCATCCGC ACCTACACCT ATGCAGACAC CCCTGATGAT     540
117
118  TTCCAGTTGC ACAACTTCAG CCTCCCAGAG GAAGATACCA AGCTCAAGAT ACCCCTGATT     600
119
120  CACCGAGCCC TGCAGTTGGC CCAGCGTCCC GTTTCACTCC TTGCCAGCCC CTGGACATCA     660
121
122  CCCACTTGGC TCAAGACCAA TGGAGCGGTG AATGGGAAGG GGTCACTCAA GGGACAGCCC     720
123
124  GGAGACATCT ACCACCAGAC CTGGGCCAGA TACTTTGTGA AGTTCCTGGA TGCCTATGCT     780
125
126  GAGCACAAGT TACAGTTCTG GGCAGTGACA GCTGAAAATG AGCCTTCTGC TGGGCTGTTG     840
127
128  AGTGGATACC CCTTCCAGTG CCTGGGCTTC ACCCCTGAAC ATCAGCGAGA CTTCATTGCC     900
129
130  CGTGACCTAG GTCCTACCCT CGCCAACAGT ACTCACCACA ATGTCCGCCT ACTCATGCTG     960
131
132  GATGACCAAC GCTTGCTGCT GCCCCACTGG GCAAAGGTGG TACTGACAGA CCCAGAAGCA    1020
133
134  GCTAAATATG TTCATGGCAT TGCTGTACAT TGGTACCTGG ACTTTCTGGC TCCAGCCAAA    1080
135
136  GCCACCCTAG GGGAGACACA CCGCCTGTTC CCCAACACCA TGCTCTTTGC CTCAGAGGCC    1140
137
138  TGTGTGGGCT CCAAGTTCTG GGAGCAGAGT GTGCGGCTAG GCTCCTGGGA TCGAGGGATG    1200
139
140  CAGTACAGCC ACAGCATCAT CACGAACCTC CTGTACCATG TGGTCGGCTG GACCGACTGG    1260
141
142  AACCTTGCCC TGAACCCCGA AGGAGGACCC AATTGGGTGC GTAACTTTGT CGACAGTCCC    1320
143
144  ATCATTGTAG ACGTCACCAG GGACACGTTT TACAAACAGC CCATGTTCTA CCACCTTGGC    1380
145
146  CACTTCAGCA AGTTCATTCC TGAGGGCTCC CAGAGAGTGG GGCTGGTTGC CAGTCAGAAG    1440
147
148  AACGACCTGG ACGCAGTGGC ACTGATGCAT CCCGATGGCT CTGCTGTTGT GGTCGTGCTA    1500
149
150  AACCGCTCCT CTAAGGATGT GCCTCTTACC ATCAAGGATC CTGCTGTGGG CTTCCTGGAG    1560
151
152  ACAATCTCAC CTGGCTACTC CATTCACACC TACCTGTGGC GTCGCCAGAA TTCGGACTAC    1620
```

AAGGACGACG ATGACAAGTT GA 1642

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 546 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Glu Phe Ser Ser Pro Ser Arg Glu Glu Cys Pro Lys Pro Leu Ser
1               5                  10                  15

Arg Val Ser Ile Met Ala Gly Ser Leu Thr Gly Leu Leu Leu Leu Gln
            20                  25                  30

Ala Val Ser Trp Ala Ser Gly Ala Arg Pro Cys Ile Pro Lys Ser Phe
        35                  40                  45

Gly Tyr Ser Ser Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser
    50                  55                  60

Phe Asp Pro Pro Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu
65                  70                  75                  80

Ser Thr Arg Ser Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln
                85                  90                  95

Ala Asn His Thr Gly Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln
            100                 105                 110

Lys Phe Gln Lys Val Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala
        115                 120                 125

Ala Leu Asn Ile Leu Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu
    130                 135                 140

Lys Ser Tyr Phe Ser Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val
145                 150                 155                 160

Pro Met Ala Ser Cys Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp
                165                 170                 175

Thr Pro Asp Asp Phe Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp
            180                 185                 190
```

```
206
207         Thr Lys Leu Lys Ile Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln
208                 195                 200                 205
209
210         Arg Pro Val Ser Leu Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu
211             210                 215                 220
212
213         Lys Thr Asn Gly Ala Val Asn Gly Lys Gly Ser Leu Lys Gly Gln Pro
214         225                 230                 235                 240
215
216         Gly Asp Ile Tyr His Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu
217                         245                 250                 255
218
219         Asp Ala Tyr Ala Glu His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu
220                     260                 265                 270
221
222         Asn Glu Pro Ser Ala Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu
223                 275                 280                 285
224
225         Gly Phe Thr Pro Glu His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly
226             290                 295                 300
227
228         Pro Thr Leu Ala Asn Ser Thr His His Asn Val Arg Leu Leu Met Leu
229         305                 310                 315                 320
230
231         Asp Asp Gln Arg Leu Leu Leu Pro His Trp Ala Lys Val Val Leu Thr
232                         325                 330                 335
233
234         Asp Pro Glu Ala Ala Lys Tyr Val His Gly Ile Ala Val His Trp Tyr
235                     340                 345                 350
236
237         Leu Asp Phe Leu Ala Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg
238                 355                 360                 365
239
240         Leu Phe Pro Asn Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser
241             370                 375                 380
242
243         Lys Phe Trp Glu Gln Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met
244         385                 390                 395                 400
245
246         Gln Tyr Ser His Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly
247                         405                 410                 415
248
249         Trp Thr Asp Trp Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp
250                     420                 425                 430
251
252         Val Arg Asn Phe Val Asp Ser Pro Ile Ile Val Asp Val Thr Lys Asp
253                 435                 440                 445
254
255         Thr Phe Tyr Lys Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys
256             450                 455                 460
257
258         Phe Ile Pro Glu Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/713,928B*

DATE: 09/01/98
TIME: 10:49:51

INPUT SET: S28333.raw

Line	Error	Original Text